AFRL-RI-RS-TR-2010-218

THEORY AND PRACTICE OF COMPRESSED SENSING IN COMMUNICATIONS AND AIRBORNE NETWORKING

STATE UNIVERSITY OF NEW YORK AT BUFFALO

DECEMBER 2010

FINAL TECHNICAL REPORT

***APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED*.**

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■**UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2010-218 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/

WALTER A. KOZIARZ
Work Unit Manager

/s/

EDWARD J. JONES, Deputy Chief
Advanced Computing Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*: December 2010

2. REPORT TYPE: Final Technical Report

3. DATES COVERED *(From - To)*: June 2009 – June 2010

4. TITLE AND SUBTITLE
THEORY AND PRACTICE OF COMPRESSED SENSING IN COMMUNICATIONS AND AIRBORNE NETWORKING

5a. CONTRACT NUMBER: FA8750-09-1-0197

5b. GRANT NUMBER: N/A

5c. PROGRAM ELEMENT NUMBER: 63662D

6. AUTHOR(S)
Stella N. Batalama

5d. PROJECT NUMBER: WCNA

5e. TASK NUMBER: BU

5f. WORK UNIT NUMBER: FF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
State University of New York at Buffalo
School of Engineering and Applied Sciences
332 Bonner Hall
Buffalo NY 14260

8. PERFORMING ORGANIZATION REPORT NUMBER: N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory/RITB
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S): AFRL/RI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER: AFRL-RI-RS-TR-2010-218

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. PA# 88ABW-2010-6397
Date Cleared: 6 December 2010.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
We consider the problem of compressed sensing and propose new deterministic constructions of compressive sampling matrices based on finite-geometry generalized polygons. For the noiseless measurements case, we develop a novel recovery algorithm for strictly sparse signals that utilizes the geometry properties of generalized polygons and exhibits complexity linear in the sparsity value. In the presence of measurement noise, recovery of the generalized-polygon sampled signals can be carried out most effectively using a belief propagation algorithm. Experimental studies included in this report illustrate our theoretical developments.

15. SUBJECT TERMS
Belief propagation, bipartite graphs, compressed sensing, compressive sampling, finite geometry, generalized polygons, Nyquist sampling, sparse matrices.

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
U	U	U	UU	17

19a. NAME OF RESPONSIBLE PERSON: WALTER A. KOZIARZ

19b. TELEPHONE NUMBER (*Include area code*): N/A

TABLE OF CONTENTS

LIST OF FIGURES

SUMMARY

We consider the problem of compressed sensing and propose new deterministic constructions of compressive sampling matrices based on finite-geometry generalized polygons. For the noiseless measurements case, we develop a novel recovery algorithm for strictly sparse signals that utilizes the geometry properties of generalized polygons and exhibits complexity linear in the sparsity value. In the presence of measurement noise, recovery of the generalized-polygon sampled signals can be carried out most effectively using a belief propagation algorithm. Experimental studies included in this report illustrate our theoretical developments.

1.0 INTRODUCTION

1.1. Report development

Compressed sensing (CS) emerges as a new technology for the acquisition of sparse signals. Most natural signals are seen to be sparse with respect to some basis or dictionary of waveforms. Compressed sensing attempts to perform direct sub-Nyquist sampling of such signals that would still allow (near-)perfect reconstruction later on as needed. Specifically, compressed sensing aims at capturing all (most) information present in a high-dimensional sparse signal using only a relatively small number of linear signal projections. The CS line of research, therefore, involves development of measurement matrices with good compression ratio coupled with effective recovery algorithms of low computation complexity.

Candes and Tao [1], and Donoho [2] demonstrated that a sparse signal can be recovered from $O(k\log(\frac{n}{k}))$ random linear measurements with high probability under certain plurality and energy preserving conditions on the random measurement matrix by solving an ℓ_1-norm minimization linear programming problem of complexity $O(n^3)$, where k, n denote the sparsity (number of non-zero elements) and dimensionality of the original signal, respectively. As the dimensionality n grows, $O(n^3)$ linear programming for ℓ_1-minimization becomes quickly impractical for implementation. An alternative iterative greedy algorithm of complexity $O(nk^2)$ termed Orthogonal Matching Pursuit (OMP) was suggested in [3]. A modified version of OMP, called Stage-wise Orthogonal Matching Pursuit (StOMP), identifies multiple non-zero components at each stage compared to only one in OMP and results in the improved complexity of $O(n\log n)$.

The work mentioned above, so far, is based on random measurement matrices where recovery can fail with non-zero probability. For hardware implementation and critical field applications, it is highly desirable to deploy an explicit well-analyzed and understood measurement matrix for compressed sensing. In [5], deterministic measurement matrices are reported for strictly sparse signals with $k = O(\sqrt{m})$ where m is the number of measurements. A matrix with $O(k\log^2 n)$measurement rows in conjunction with a recovery algorithm of complexity $O(k\log^2 n)$ is developed in [6]. In [7], adjacency matrices from expander graphs are used as measurement matrices with a recovery algorithm of complexity linear in n. Criteria to use the suitability of deterministic measurement matrices to operate on most k-sparse signals are provided in [8].

In this work, we construct new deterministic sparse measurement matrices in $\{0, 1\}^{m\times n}$ for compressed sensing from classical finite-geometry generalized polygons [9], [10], [11], [12]. For noiseless sensing, we develop a novel algorithm that iteratively recovers the original k-sparse signal in k iterations by exploiting the property that the generalized-polygon measured vector involves only structured sums of subsets of the elements of the original signal. The proposed algorithm is applicable to strictly k-sparse input signals. Extensions to the case where the input signals are less than k-sparse and/or are sampled in the presence of noise are also considered. Motivated by the success of decoding generalized-polygon low-density parity-check (LDPC) codes, we exploit the iterative belief propagation (BP) algorithm as the means to efficiently recover the original signals from compressed measurements, thus deviating for both noiseless and noisy sampling from conventional ℓ_1-norm recovery procedures.

2.0 METHODS, ASSUMPTIONS, PROCEDURES, RESULTS AND DISCUSSION

2.1. Compressed sensing background and problem statement

We consider a family of signals $\mathbf{x}\in R^n$ that are k-sparse if all but k elements of $\mathbf{x}$ are zero. The sparse signals can also be extended to the set of signals that have a sparse representation in some particular basis. For simplicity of our presentation, we assume that all sparse signals are represented in a standard basis (the $n\times n$ identity matrix). In compressed sensing, it is unnecessary to measure all the n elements of the signal $\mathbf{x}$; instead, we only need to measure a small number of linear combinations of elements of $\mathbf{x}$, i.e.

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1}$$

where $\mathbf{A}$ denotes an $m\times n$ linear measurement matrix with $m \ll n$ and $\mathbf{x}$ is a k-sparse signal in $\mathbb{R}^n$, $k<m$. Most existing work is focusing on probabilistic constructions where entries of $\mathbf{A}$ are drawn from independent and identically distributed (i.i.d.) Gaussian or Bernoulli random variables. With these measurement matrices, it is shown [1] that given the number of measurements $m = O(k\log(\frac{n}{k}))$, the k-sparse signal $\mathbf{x}$ can be exactly recovered via ℓ_1 optimization

$$\hat{\mathrm{x}} = arg\min_{\tilde{\mathrm{x}}} \|\tilde{\mathrm{x}}\|_{\ell_1} \qquad \text{subject to} \quad \mathbf{y} = \mathbf{A}\tilde{\mathrm{x}} \tag{2}$$

with overwhelming probability. Nevertheless, when the dimensionality n is large, this scheme has two significant drawbacks: (i) Random measurement matrices require large storage space; (ii) the ℓ_1 minimization exhibits complexity $O(n^3)$ which is too high for implementation. It is highly desirable to develop deterministic constructions with low storage requirements and low complexity recovery. In the following section, we present novel deterministic sparse measurement matrices associated with efficient recovery algorithms for compressed sensing.

2.2. GP-based measurement matrices and recovery algorithms

2.2.1. Measurement matrices based on generalized polygons

A finite-geometry generalized polygon is referred to as an incidence structure $\Gamma = (\mathfrak{P}, \mathfrak{L}, \mathfrak{I})$ including the point set $\mathfrak{P}$, the line set $\mathfrak{L}$, and the incidence set $\mathfrak{I} \subseteq \mathfrak{P} \times \mathfrak{L}$. A point $p \in \mathfrak{P}$ is incident to a line $l \in \mathfrak{L}$, say $p\mathfrak{I}l$, if and only if $(p, l) \in \mathfrak{I}$. Thus, the existence of incidence between points and lines can be represented by undirected bipartite graph G, as illustrated in Fig.1. We say that Γ is a weak generalized polygon if and only if the diameter of G is half of its girth. Γ is a generalized polygon if and only if it is a weak generalized polygon and each vertex of G has degree at least three. If all point vertices have the same degree $t + 1$ and all line vertices have the same degree $s+1$, then G is regular and Γ has order (s, t). The incidence matrix of Γ is the matrix with rows labeled by the lines of Γ, columns labeled by the points of Γ, and entries 1 or 0 depending on the corresponding point-to-line incidence. It is straightforward to use the incidence matrix for a given generalized polygon Γ as the measurement matrix for compressed

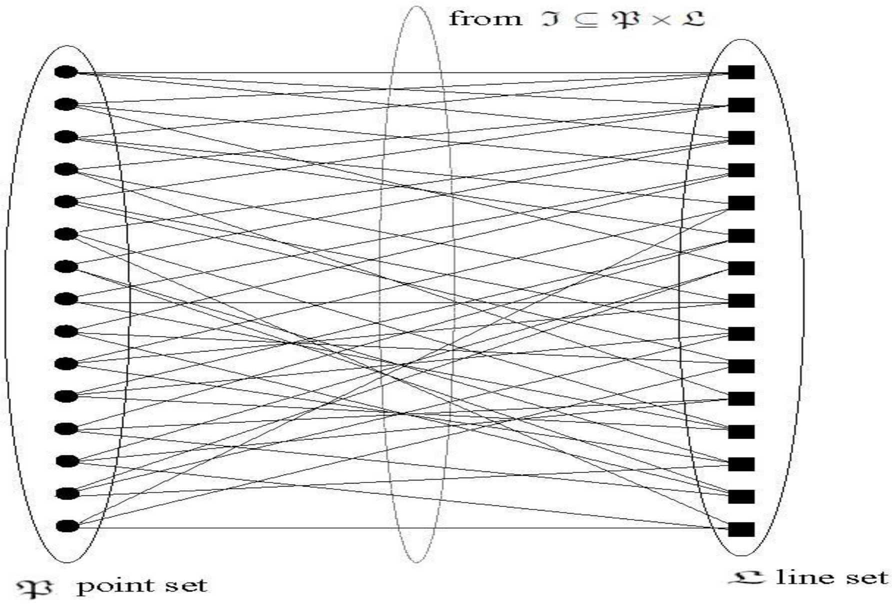

Figure 1: Bipartite graph representation of a generalized polygon Γ.

sensing. Among all known generalized polygons [12], two families of generalized polygons, hermitian generalized quadrangles $H(3, q^2)$ and twisted triality hexagon $T(q^3, q)$, lead to the measurement matrix with an excellent compression ratio. In detail, the measurement matrix from $H(3, q^2)$ is of size $(q^3+ 1)(q + 1) \times (q^3+ 1)(q^2+ 1)$ corresponding to a compression ratio $(q+1) : (q^2+1)$ while the one from $T(q^3,q)$ is of size $(q^8+q^4+1)(q+1)\times(q^8+q^4+1)(q^3+1)$ corresponding to a compression ratio $(q+1) : (q^3+ 1)$, $q \geq 2$ and $q \in \mathbb{Z}$. On the other hand, the storage requirement is also a key concern for compressed sensing construction. The incidence structure for $H(3, q^2)$ and $T(q^3, q)$ has order (q^2, q) and (q^3, q), respectively. Thus, the measurement matrix from $H(3, q^2)$ and $T(q^3, q)$ requires lower storage ($O(q^6)$ and $O(q^{12})$, respectively) than the storage requirements of random measurement matrices, ($O(q^9)$ and $O(q^{20})$, correspondingly).

2.2.2. Recovery algorithm for the noiseless case

We collect compressed sensing observations from a GP-based measurement matrix **A** in the noiseless environment, i.e.

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \tag{3}$$

where the entries of **A** are restricted to {0, 1}. The measurements are seen to be sums of small subsets of the coordinates of **x**. Let $S \in supp(\mathbf{x})$ be the support set of **x**, which is defined as the set of indices corresponding to the non-zero coordinates of **x**. We assume that for any two subsets of S, $S_1, S_2 \subseteq S$ with $S_1 \neq S_2$, the following condition holds:

$$\sum_{i\in S_1} x_i \neq \sum_{i\in S_2} x_i. \tag{4}$$

The measurement matrix **A** based on the geometry structure of generalized polygons has the following properties. The proofs of Propositions 2 and 3 are given in Section 2.4.

Proposition 1 *For any two rows (columns) of* **A***, there exists at most one column (row) with non-zero entries in common.*

Iterative recovery algorithm

Initialize $\widehat{\mathrm{x}}^{(0)} = 0$ and $\mathrm{y}^{(0)} = \mathrm{y}$.

For $t = 1, 2, \ldots, k$

(i) Identify the index c such that at least two of the measurements corresponding to $supp(\mathrm{a}_c)$ take a non-zero identical value, say d.

(ii) Update $\widehat{\mathrm{x}}^{(t)} = \widehat{\mathrm{x}}^{(t-1)} + d\mathrm{e}_c$ and $\mathrm{y}^{(t)} = \mathrm{y}^{(t-1)} - d\mathrm{a}_c$.

End

Output: Recovery signal $\widehat{\mathrm{x}}^{(k)}$.

Figure 2: Proposed iterative recovery algorithm in the noiseless environment.

Proof : Rows and columns stand for lines and points in generalized polygons, respectively. Any two different lines have at most one intersection point while any two different points determine one line.

Proposition 2 *Given $1 \leq u \leq 2q - 1$, the submatrix* $\mathbf{A}(u)$ *that consists of any u columns of the matrix* $\mathbf{A}$ *generated from $H(3, q^2)$ has at least two rows of weight* 1 *located in the same column.*

Proposition 3 *Given $1 \leq u \leq 3q - 2$, the submatrix* $\mathbf{A}(u)$ *that consists of of any u columns of the matrix* $\mathbf{A}$ *generated from $T(q^3, q)$ has at least two rows of weight* 1 *located in the same column.*

We are now ready to proceed by an iterative recovery algorithm. We first initialize at $\hat{\mathbf{x}}^{(0)} = \mathbf{0}$ and $\mathbf{y}^{(0)} = \mathbf{y}$ where superscripts denote iteration index. Let $\mathbf{a}_i$ denote the i-th column of the measurement matrix $\mathbf{A}$. At the first iteration, Proposition 2 (or Proposition 3) implies that under the assumption in (4) and the condition $k \leq 2q - 1$ (or $3q - 2$) for $\mathbf{A}$ generated from $H(3, q^2)$ (or T (q^3, q)), respectively, we can identify the index c, $1 \leq c \leq n$, such that the measurements of $\mathbf{y}^{(0)}$ corresponding to supp($\mathbf{a}_c$) have at least two of them taking a non-zero identical value, say d. We determine the c-th element of $\mathbf{x}$, and compute the estimate update by $\hat{\mathbf{x}}^{(1)} = \hat{\mathbf{x}}^{(0)} + d\mathbf{e}_c$, where $\mathbf{e}_c$ denotes the unit-length vector in $\mathbb{R}^n$ whose elements are all zero except for the c-th element which is equal to one. Furthermore, we also update the measurement vector by $\mathbf{y}^{(1)} = \mathbf{y}^{(0)} - d\mathbf{a}_c$. Thus, the sparsity is decreasing to $k - 1$. It is clear that Propositions 2 and 3 still hold with the sparsity $\leq k - 1$. Repeating the procedure $k - 1$ times/iterations, we can finally obtain the estimate $\hat{\mathbf{x}}^{(k)}$ that is an exact recovery of $\mathbf{x}$. The iterative algorithm is summarized in Fig. 2. We note that the above procedure terminates after k iterations, and each iteration requires only two vector updates. The overall complexity of this algorithm is $O(k)$.

2.2.3. Recovery algorithm for the noisy case

The recovery algorithm above is confined to the noiseless case for the strictly sparse input signal. In fact, most compressed sensing systems deal with moderately sparse input signals in presence of noise. Motivated by LDPC decoding, we exploit a belief propagation algorithm for compressed sensing recovery. A BP algorithm [13] can efficiently compute the marginal posterior distribution of all variables conditioned on the observations by iteratively exchanging messages over the edges of the bipartite graph. First, we express the compressed sensing measurements corrupted by noise as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \qquad (5)$$

where $\mathbf{n}$ denotes a zero-mean additive white Gaussian noise (AWGN) with covariance matrix $\sigma^2\mathbf{I}$. To ensure signal sparseness, we impose a two-state mixture prior on $\mathbf{x}$. Let the i-th element of $\mathbf{x}$ be $x_i = B_i G_i$, where the state variable B_i is Bernoulli with $P_r\{B_i = 1\} = \frac{k}{n}$ and $P_r\{B_i = 0\} = \frac{n-k}{n}$, and G_i is Gaussian with mean zero and large variance σ^2_s . At the state $B_i = 1$, the element x_i takes a non-zero large value drawn from the large-variance Gaussian distribution. On the other hand, x_i is set to 0 for the state $B_i = 0$. The overall pdf $f(x_i)$ of this distribution is illustrated in Fig. 3. Moreover, it is assumed that the coordinates $x_1, x_2, \ldots, x_n$ are mutually independent. The two-state mixture model makes the signal $\mathbf{x}$ k-sparse as n is large.

We consider the bipartite graph representation of $\mathbf{A}$ in Fig. 1. In the graph, the message along the edge is defined as the pdf of marginal distribution of the point variable associated with the edge. Let $M^{(t)}_{p\to l}$ denote the message sent from the point node p to the line node l at t-th iteration, and let $M^{(t)}_{l\to p}$ denote the message coming out of the line node l to the point node p at t-th iteration. The BP algorithm starts with messages $M^{(0)}_{p\to l}$ from point to line nodes ($f(x_i)$). In each iteration of BP, messages first propagate over the edges from line nodes to point nodes. Subsequently, each point node generates the messages based on the previously received messages, and sends them back to the line nodes. The messages from point nodes are used to compute the new messages for the next iteration. The message computation is conforming to the following rules [14]:

(i) The message $M^{(t)}_{p\to l}$ is the product of all the messages received at p except the one coming

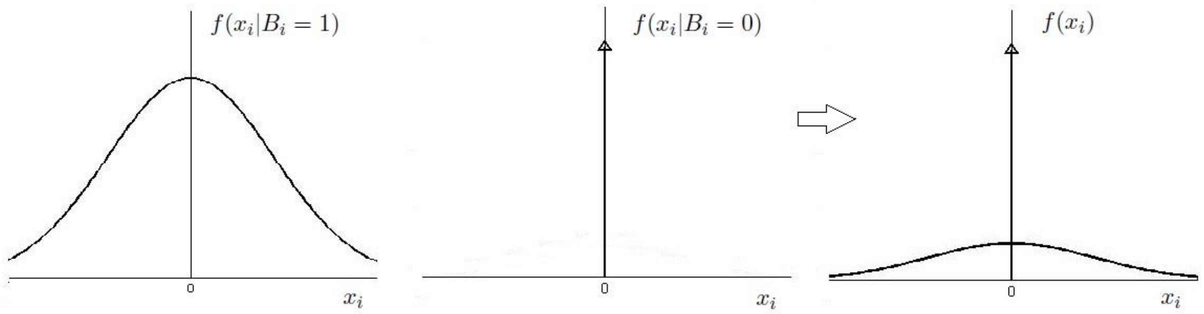

Figure 3: The overall pdf $f(x_i)$ of a two-state mixture distribution.

out of l, i.e.

$$M_{p\to l}^{(t)}(x_p) \propto \prod_{g\in N(p)-\{l\}} M_{g\to p}^{(t-1)}(x_p), \tag{6}$$

where the operator "∝" means that $M_{p\to l}^{(t)} = c\prod_{g\in N(p)-\{l\}} M_{g\to p}^{(t-1)}$ for some constant c such that $M_{p\to l}^{(t)}$ is a probability function, and $N(p) - \{l\}$ is the set of neighbors of p excluding l.

(ii) The message $M_{l\to p}^{(t)}$s updated by the convolution for all neighbors of l except p, i.e.

$$M_{l\to p}^{(t)}(x_p) \propto \int_{\sim\{p\}} p(y_l|\mathbf{x}) \prod_{g\in N(l)-\{p\}} M_{g\to l}^{(t)}(x_g) d(\sim\{p\}), \tag{7}$$

where $\sim\{p\}$ is the set of neighbors of l excluding p, and $p(y_l|\mathbf{x})$ is the conditional probability density function of y_l given $\mathbf{x}$.
After several iterations, the marginal pdf for a random variable $\mathbf{x}_p$ is estimated by the product of all most recent incoming messages along the edges connecting to that node:

$$f(x_p) \propto \prod_{l\in N(p)} M_{l\to p}^{(T)}(x_p). \tag{8}$$

Based on the marginal distribution, we can extract MAP estimates for elements of $\mathbf{x}$. If the bipartite graph is free of cycles, the BP algorithm provides an exact marginal posterior distribution. However, for the GP-based graphs with few loops, it only offers an approximate inference [15].

2.3. Simulation studies

In this section, we investigate the performance of compressed sensing systems with GP-based measurement matrices and their recovery algorithms. We consider the sparse input signal $\mathbf{x}$ with sparsity k. The positions of k non-zero coordinates in $\mathbf{x}$ are picked uniformly at random

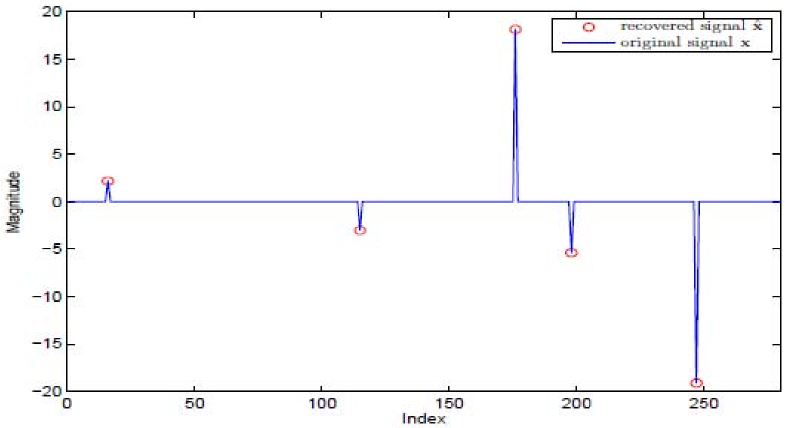

Figure 4: Original signal $\mathbf{x}$ and recovered signal $\hat{\mathbf{x}}$ using the iterative recovery algorithm in Section 2.2.2. in the noiseless environment.

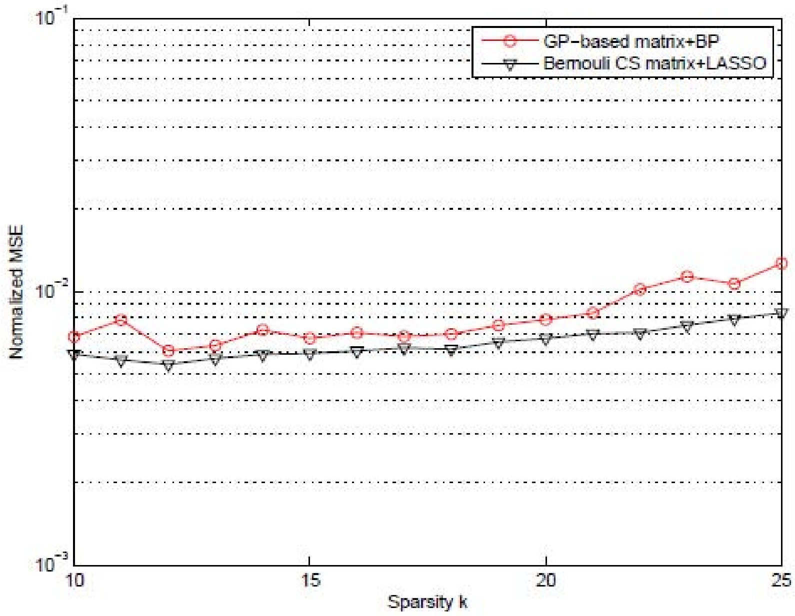

Figure 5: Normalized MSE as a function of the sparsity k using the measurement matrix **A** from $H(3, q^2)$ and BP recovery algorithm.

and their values are drawn from the Gaussian distribution with mean $\mu = 0$ and standard deviation $\sigma_S = 10$. In the first example, we are dealing with compressed sensing in the noiseless environment. The measurement matrix **A** is generated from the generalized quadrangle $H(3, 3^2)$ with $m = 112$ and $n = 280$ as described in Section 2.2.1., and the sparsity k of **x** is set to be 5. From Fig. 4, we observe that the iterative algorithm in Section 2.2.2 performs an exact recovery when the input signal **x** is strictly sparse.

In the next study, we compare our scheme of the GP-based measurement matrix and the BP recovery algorithm with the conventional scheme of random Bernoulli measurement matrix associated with the LASSO algorithm [16] in the noisy environment. We still use the incidence matrix of $H(3, 3^2)$ as the measurement matrix **A** and the same signal **x**. The maximum iteration number of the BP algorithm is set equal to 10. The total 1000 experiments are carried out in the presence of noise that is taken to be Gaussian with mean 0 and covariance matrix **I**. In Fig. 5, we plot as a function of the sparsity k the normalized mean-square-error (MSE) that is defined as

$$\text{normalized}\, MSE = E\{\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|^2}{\|\mathbf{x}\|^2}\}. \tag{9}$$

For effective visualization, we use a log scale for the y-axis. We observe that both schemes perform similarly when the sparsity k varies from 10 to 21. The conventional scheme performs slightly better than the proposed scheme after $k = 21$. However, the LASSO algorithm exhibits the complexity of convex programming, $O(n^3)$. In the contrast, the overall complexity of the BP algorithm is $O(n \log^2 n)$ [17]. In Fig. 6, we take the average normalized ℓ_1-error, i.e. $E\{\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_1}{\|\mathbf{x}\|_1}\}$, as a performance metric. We observe that the proposed scheme compares favorably to the conventional scheme.

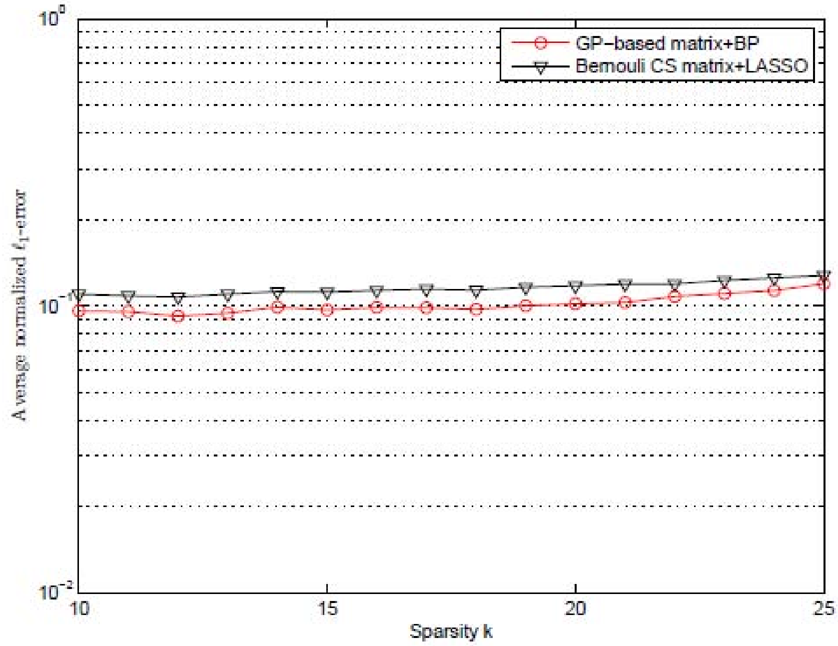

Figure 6: Average normalized ℓ_1-error as a function of the sparsity k using the measurement matrix **A** from $H(3, q^2)$ and BP recovery algorithm.

2.4. Proof of propositions

Proof of Proposition 2

For $u = 1$, it is straightforward to find two rows of $\mathbf{A}(u)$ that have weight 1 in the same column. In the following, we are focusing on $2 \le u \le 2q - 1$. First, the submatrix $\mathbf{A}(u)$ falls into one of two cases.

Case 1: There exists no row of weight 2 or more in $\mathbf{A}(u)$. Of course, $\mathbf{A}(u)$ has at least two rows of weight 1 in the same column.

Case 2: There exists at least one row of weight 2 or more in $\mathbf{A}(u)$. Proposition 1 implies that two columns of $\mathbf{A}(u)$, say, the i-th and j-th columns, have the property that the intersection set of supp($\mathbf{a}_i$) and supp($\mathbf{a}_j$) contains only one element, say r, $1 \le r \le m$. Because each of $\mathbf{a}_i$ and $\mathbf{a}_j$ has $q+1$ non-zero entries and only the r-th row of $\mathbf{A}(u)$ has two 1's at the i-th and j-th coordinates, a total of $2q + 1$ rows of A(u) are incident to two columns $\mathbf{a}_i$ and $\mathbf{a}_j$. Let $\mathbf{A}(u, i, j)$ be the submatrix formed by the $2q + 1$ rows of $\mathbf{A}(u)$. Without loss of generality, let $\mathbf{a}_s$ be one column of $\mathbf{A}(u)$ different than $\mathbf{a}_i$ and $\mathbf{a}_j$. The columns $\mathbf{a}_i$, $\mathbf{a}_j$ and $\mathbf{a}_s$ correspond to the points p_i , p_j and p_s, respectively If the point p_s is incident to the line l_r which both points p_i and p_j are incident to, then the column of the matrix $\mathbf{A}(u, i, j)$ corresponding to p_s has only one nonzero entry by Proposition 1. Otherwise, p_s is collinear with either of p_i and p_j or none of them by the following lemma [18].

Lemma 1 *For every non-incident pair (p, l) of a generalized quadrangle, there exists a unique pair* $(p', l') \in \mathfrak{P} \times \mathfrak{L}$ *, for which* $p \,\mathfrak{I}\, l' \,\mathfrak{I}\, p' \,\mathfrak{I}\, l$.

This implies that there exists at most one nonzero entry on each column of $\mathbf{A}(u, i, j)$ except for the two columns corresponding to p_i and p_j.

Now, we count the total number of nonzero entries in $\mathbf{A}(u, i, j)$ which is at most $2(q+1)+u-2$. Suppose that v is the number of rows in $\mathbf{A}(u, i, j)$ with weight 1. Then, $2q + 1 - v$ row has at least weight 2. We can obtain the inequality

$$v + 2(2q + 1 - v) \leq 2(q + 1) + u - 2. \tag{10}$$

For $2 \leq u \leq 2q - 1$, the inequality (10) leads to

$$v \geq 2q + 2 - u \geq 3. \tag{11}$$

At least two indices of rows with weight 1 belong to either supp($\mathbf{a}_i$) or supp($\mathbf{a}_j$) which completes the proof.

Proof of Proposition 3

Proposition 3 can be proved in the same manner as Proposition 2. The case of u = 1, 2 is straightforward and thus omitted. In the following, we focus on $3 \leq u \leq 3q - 2$. We consider three cases for $\mathbf{A}(u)$: (i) No row of weight 2 or more; (ii) only one row of weight 2 and no row of weight more than 2; (iii) two or more rows of weight 2 or more. For the first two cases, it is easy to find two rows of weight 1 in the same column. For the last case, we select three different columns of $\mathbf{A}(u)$, say h-th, i-th and j-th columns denoted by $\mathbf{a}_h$, $\mathbf{a}_i$ and $\mathbf{a}_j$, having the following property

$$|supp(\mathbf{a}_h) \cup supp(\mathbf{a}_i) \cup supp(\mathbf{a}_j)| = 3q + 1. \tag{12}$$

Let $\mathbf{A}(u, h, i, j)$ be the submatrix formed by the $3q + 1$ rows of $\mathbf{A}(u)$. The points p_h, p_i, p_j that correspond to $\mathbf{a}_h$, $\mathbf{a}_i$, $\mathbf{a}_j$ exhibit one of the following geometric relations: (1) All three points are incident to the same line; (2) p_i , p_j are collinear with p_h, but p_i is not collinear with p_j; (3) p_h is collinear with p_i , p_i is collinear with p_j , and p_j is not collinear with p_h by the following lemma in [18].

Lemma 2 *For every non-incident pair (p, l) of a generalized hexagon, there exists a unique pair* $(p', l') \in \mathfrak{P} \times \mathfrak{L}$ *for which* $p \,\mathfrak{I}\, l' \,\mathfrak{I}\, p' \,\mathfrak{I}\, l$ *, or a unique quadruple* $(p', l', p'', l'') \in \mathfrak{P} \times \mathfrak{L} \times \mathfrak{P} \times \mathfrak{L}$ *for which* $p \,\mathfrak{I}\, l'' \,\mathfrak{I}\, p'' \,\mathfrak{I}\, l' \,\mathfrak{I}\, p' \,\mathfrak{I}\, l$.

For all three scenarios, Lemma 2 also implies that there exists at most one non-zero entry on each column of $\mathbf{A}(u, h, i, j)$ except the three columns corresponding to p_h, p_i, p_j.

We count the total number of nonzero entries in $\mathbf{A}(u, h, i, j)$ which is at most $3(q+1)+u-3$. Suppose that v is the number of rows in $\mathbf{A}(u, h, i, j)$ with weight 1. Then, $3q + 1 - v$ row has at least weight 2. We can obtain the inequality

$$v + 2(3q + 1 - v) \leq 3(q + 1) + u - 3. \tag{13}$$

For $3 \leq u \leq 3q - 2$, the inequality (13) leads to

$$v \geq 3q + 2 - u \geq 4. \tag{14}$$

This implies that at least two indices of rows with weight 1 belong to one of $supp(\mathbf{a}_h)$, $supp(\mathbf{a}_i)$ and $supp(\mathbf{a}_j)$.

3.0 CONCLUDING REMARKS

In this work, we constructed two families of deterministic sparse measurement matrices generated from finite-geometry generalized polygons, H(3, q^2) and T (q^3, q). Exploiting the geometry feature of GPs, we develop an iterative recovery algorithm that achieves perfect recovery for strictly sparse signals in the noiseless environment. This algorithm is highly efficient and requires linear computation complexity $O(k)$. We also considered compressed sensing systems with moderately sparse input signals in the presence of noise. We developed a belief propagation algorithm for compressed sensing recovery that exhibits nearly linear complexity $O(n \log^2 n)$. Simulation studies illustrated our theoretical developments.

References

[1] E. Candes and T. Tao, “Near optimal signal recovery from random projections: Universal encoding strategies?” *IEEE Trans. Inform. Theory*, vol. 52, pp. 5406-5425, Dec. 2006.

[2] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.

[3] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inform. Theory*, vol. 53, pp. 4655-4666, Dec. 2007.

[4] D. L. Donoho, Y. Tsaig, I. Drori, and J. C. Starck, “Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit,” Stanford Statistics Dept., Stanford Univ., Stanford, CA, TR-2006-2, Mar. 2006.

[5] R. A. DeVore, “Deterministic constructions of compressed sensing matrices,” *J. Complexity*, vol. 23, pp. 918-925, Aug. 2007.

[6] G. Cormode and S. Muthukrishnan, “Combinatorial algorithms for compressed sensing,” *Lecture Notes in Compu. Sci.*, vol. 4056, pp. 280-294, 2006.

[7] W. Xu and B. Hassibi, “Efficient compressive sensing with deterministic guarantees using expander graphs,” in *Proc. IEEE Inform. Theory Workshop*, Lake Tahoe, CA, Sep. 2007, pp. 414-419.

[8] R. Calderbank, S. Howard, and S. Jafarpour, “Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property,” *IEEE J. Select. Topics Signal Proc.*, vol. 4, pp. 358-374, Apr. 2010.

[9] J. Tits, “Sur la Trialite et Certains Groupes qui s’en Deduisent,” *Inst. Hautes Etudes Sci. Publ. Math.*, vol. 2, pp. 14-60, 1959.

[10] W. Feit and G. Higman, “The nonexistence of certain generalized polygons,” *J. Algebra*, vol. 1, pp. 114-131, 1964.

[11] H. van Maldeghem, *Generalized Polygons*. Basel, Switzerland: Birkhauser-Verlag, 1998.

[12] Z. Liu and D. A. Pados, “LDPC codes from generalized polygons,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 3890-3898, Nov. 2005.

[13] R. G. Gallager, *Low-density Parity-check Codes*. Cambridge, MA: MIT Press, 1963.

[14] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory,* vol. 47, pp. 498-519, Feb. 2001.

[15] J. Mooij and H. Kappen, "Sufficient conditions for convergence of the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 53, pp. 4422-4437, Dec. 2007.

[16] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Math*., vol. 59, pp. 1207-1223, Aug. 2006.

[17] D. Baron, S. Sarvotham, and R. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Trans. Signal Proc*., vol. 58, pp. 269-280, Jan. 2010.

[18] H. V. Maldeghem, "A finite generalized hexagon admitting a group acting transitively on ordered heptagons is classical," *J. Combin. Theory*, vol. 75, pp. 254-269, Aug. 1996.

ACRONYMS

CS: Compressed sensing
OMP: Orthogonal matching pursuit
StOMP: Stage-wise orthogonal matching pursuit
LDPC: Low-density parity-check
BP: Belief propagation
GP: Generalized polygon
AWGN: Additive white Gaussian noise
MAP: Maximum a posteriori
MSE: Mean squared error